# AN ANALYSIS OF SEQUENTIAL SAMPLING STRATEGY IN PEST CONTROL BASED ON NEGATIVE BINOMIAL DISTRIBUTION

## I. C. Kipchirchir

School of Mathematics
University of Nairobi
Nairobi, Kenya

Correspondence: School of Mathematics, University of Nairobi, P.O. Box 30197-00100, Nairobi. Kenya.    Email: kipchirchir@uonbi.ac.ke

# AN ANALYSIS OF SEQUENTIAL SAMPLING STRATEGY IN PEST CONTROL BASED ON NEGATIVE BINOMIAL DISTRIBUTION

## I. C. Kipchirchir

School of Mathematics
University of Nairobi
Nairobi, Kenya

## ABSTRACT

In this paper, we analyse sequential sampling strategy based on the negative binomial distribution in the context of chemical pest control with reference to environmental pollution and food security. Type 1 error means spraying the crops when it is not necessary and type II error means not spraying the crops when it is necessary. The analysis demonstrated that keeping both probabilities of type I and II errors at possible minima provide better protection against wrong decisions, however, the sequential process requires more observations (large sample size) and more so at the state of profound indifference. On the other hand, if we want to spray when necessary, that is, when the infestation has reached economic threshold, we should keep probability of type II error at a possible minimum and if we want to avoid unnecessary pollution of the environment, we should keep probability of type I error at a possible minimum and have better protection against wrong decisions in each case without taking more observations.

**KEYWORDS:** Sequential sampling, chemical pest control, environmental pollution, food security, economic threshold, negative binomial distribution

## INTRODUCTION

Pest control consists in preventing the population of any pest species from rising to a level at which it does appreciable economic damage and is one of the most important branches of applied ecology.

We shall suppose that a farmer has realized that the crops in his farm are infested with pests. As it is normally impossible to count all the pests in a habitat, it is necessary to estimate the population by sampling. Sampling must provide accurate assessment of the state of infestation. Depending on the crop and the pest, this state is generally defined either as the number of individuals per sampling unit, or as the proportion of infested units.

To establish infestation levels, sampling has to be carried out to determine pest density. To avoid unnecessary sampling, sequential sampling is employed. A sequential plan usually allows one to stop sampling as soon as sufficient information has been gathered.

Correspondence: School of Mathematics, University of Nairobi, P.O. Box 30197-00100, Nairobi. Kenya.
Email: kipchirchir@uonbi.ac.ke

There is no universal sampling method for pest populations. The sampling of a particular pest population must be resolved about the distribution and the lifecycle of the pest involved. Assuming that the life-cycle and the stage for sampling are known, preliminary work will be necessary to gain some knowledge of the dispersion of the pest. Dispersion is the description of the pattern of distribution of organisms in space (Southwood, 1966) and is often referred to as spatial distribution.

The farmer has chemical control facilities, which he can use to control the infestation. However, because of environmental pollution, chemical control measures should not be applied unless infestation has reached economic threshold (state of infestation of a particular crop by a particular pest at which intervention to control that pest becomes economically justified).

Once the density levels that are permissible and those that we associate with extensive damage have been established, a sequential plan can be employed. We shall hypothesize that the farmer should not spray the crops if the mean number of pests per unit $m$ is less than or equal to $m_0$ and spray the crops if $m$ is greater than or equal to $m_1$ (economic threshold level).

Then, for $m_0 < m_1$, we have the two hypotheses

$$H_0 : m \leq m_0$$
$$versus\ H_1 : m \geq m_1 . \tag{1}$$

$H_0$ is the null hypothesis and $H_1$ is the alternative hypothesis. If we accept $H_0$, the farmer should not spray the crops, whereas if we reject $H_0$, the farmer should spray the crops.

Type I error means spraying the crops when it is not necessary and type II error means not spraying the crops when it is necessary. Incurring type I error results in unnecessary pollution of the environment. Incurring type II error results in crop failure leading to food insecurity.

Having protection against type I error leads to the farmer not spraying the crops when it is not necessary and having protection against type II error leads to the farmer spraying the crops when it is necessary, that is, when the infestation has reached economic threshold.

Furthermore, we are content to have

$$P(type\ I\ error) = P(rejecting\ H_0\ /H_0\ is\ true) = \alpha$$
$$and\quad P(type\ II\ error) = P(accepting\ H_0\ /H_0\ is\ false) = \beta. \tag{2}$$

Alternatively, $\alpha$ is the level of significance of the test and $1 - \beta$ is the power of the test to detect a false null hypothesis. The smaller the values of $\alpha$ and $\beta$, the better the protection against wrong decisions.

A desirable test of a statistical hypothesis is that which minimizes both $\alpha$ and $\beta$. However, it is not possible to minimize both $\alpha$ and $\beta$ simultaneously because they are so much related that a desirable decrease in one is accompanied by undesirable increase in the other. In practice, the sample size helps determine how small these probabilities may become.

## 1. SEQUENTIAL SAMPLING BASED ON NBD

The negative binomial distribution (NBD) with parameters $k$ and $p$ is defined as

$$P(X = x) = \binom{k + x - 1}{x}\left(\frac{p}{q}\right)^x \left(\frac{1}{q}\right)^k, \quad x = 0,1,2, \dots; \quad k > 0, p > 0, q = 1 + p \tag{3}$$

where $X$ is the number of pests per unit having mean $m = kp$ and variance $v = kpq$.

Anscombe (1949) gave statistical analysis of insect counts based on the negative binomial distribution. Small values of $k(k \to 0)$ are associated with overdispersed (clustered or aggregated or clumped or patchy) population whereas large values of $k(k \to \infty)$ are associated with random population (equivalently described by Poisson distribution). Moreover, negative binomial is a versatile distribution in describing dispersion. Kipchirchir (2011) demonstrated analytically the versatility of the negative binomial distribution in describing dispersion.

Oakland (1950) applied sequential analysis to white fish infested by cysts of a tapeworm. The fish were classified as large, medium and small. Fishes were dissected and cysts counted. A sequential plan was applied to the different classes. Examination of medium white fish was used for illustration and the distribution of cysts per fish was found to fit the negative binomial distribution. The formulae for a sequential plan developed by Wald (1947) were used in determination of acceptance and rejection regions. The probability of accepting a lot of fish for any possible value of incoming quality and the amount of inspection required for the test are presented.

Ba-Angood and Stewart (1980) employed a sequential sampling plan to monitor the economic thresholds of cereal aphids in South-western Quebec. Weekly counts of aphids on spring grown barley were carried out with economic threshold of 16 aphids per tiller and acceptable aphid infestation of 5 aphids per tiller. Out of 14 samples of aphid counts, only 4 fitted the Poisson distribution and the remaining fitted the negative binomial distribution which indicated that the aphids were likely to be overdispersed. The equations for decision lines pertaining to the negative binomial were used. The operating characteristic (OC) and the average sample number (ASN) functions were determined according to Oakland (1950).

Assuming the dispersion of pests is governed by the negative binomial distribution with parameter $k$, which is the same under $H_0$ and $H_1$, we have for, $p_0 < p_1$

$$H_0 : kp \le kp_0 \Leftrightarrow p \le p_0$$
$$\textit{versus } H_1 : kp \ge kp_1 \Leftrightarrow p \ge p_1. \tag{4}$$

Consider the parameter space $\Omega = \{p/p > 0\}$ and denote by $\Omega_0$ the set of all these parameter points $p$ which are consistent with $H_0$, that is, $H_0$ is precisely the statement that the true parameter point is included in the set $\Omega_0$.

The zone of preference for acceptance of $H_0$ will always be a subset of $\Omega_0$ and the zone of preference for rejection of $H_0$ will be a subset of $\bar{\Omega}_0$, that is, the true parameter point $p$ is not in the set $\Omega_0$. The zone of indifference will always consist of points of $\Omega_0$ and $\bar{\Omega}_0$ which are near the boundary or on the boundary of $\Omega_0$.

As far as controlling the crop damaging pests is concerned, for any $p \leq p_0$, spraying of crops is an error of practical importance; for any $p \geq p_1$, not spraying the crops is an error of practical importance; whereas for any value $p \in (p_0, p_1)$ there is no strong preference for either action. The zone of indifference is defined as the interval $p_0 < p < p_1$, the zone of preference for acceptance of $H_0$ as the set consisting of all values $p \leq p_0$ and the zone of preference for rejection of $H_0$ (accepting $H_1$) as the set of all values $p \geq p_1$.

Now suppose that the hypothesis $H_0$ to be tested states that the true parameter point $p$ lies in a given set $\Omega_0$ of parameter points. Then, we wish to make a probability of accepting $H_0$ as high as possible when $p$ lies in $\Omega_0$, that is,

$$Max(1 - \alpha)$$
$$p \in \Omega_0 \tag{5}$$

and as low as possible when p lies outside $\Omega_0$, that is,

$$Min(1 - \alpha)$$
$$p \in \bar{\Omega}_0 \tag{6}$$

The division of the parameter space into three zones consequently leads to the division of the sample space into three zones. Using sequential analysis procedure, two decision lines can be determined which divides the sample space into three zones (Wald, 1947). According to Wald (1947), the probability that the sequential process will terminate eventually is one.

According to Oakland (1950), decision $D(y)$, is made according to

$$D(y) = \begin{cases} reject\ H_0, & if\ \ y > sn + c_1 \\ continue\ sampling, & if\ \ sn + c_0 < y < sn + c_1 \\ accept\ H_0, & if\ \ y < sn + c_0 \end{cases} \tag{7}$$

where $y = \sum_{i=1}^{n} x_i$, $n$ is the number of units examined (sampled), the intercepts $c_0$ and $c_1$ and slope $s$ of the decision lines are given by

$$c_0 = \frac{\log\left(\frac{\beta}{1-\alpha}\right)}{\log\left(\frac{p_1 q_0}{p_0 q_1}\right)}, \quad c_1 = \frac{\log\left(\frac{1-\beta}{\alpha}\right)}{\log\left(\frac{p_1 q_0}{p_0 q_1}\right)} \quad and \quad s = \frac{k \log\left(\frac{q_1}{q_0}\right)}{\log\left(\frac{p_1 q_0}{p_0 q_1}\right)} \tag{8}$$

## 2. THE OC AND THE ASN FUNCTIONS

The Operating Characteristic (OC) function describes how well the test procedure achieves its objective of making correct decisions. The OC function $L(p)$ is defined as the probability that the sequential process will terminate with the acceptance of $H_0$ when $p$ is the true value of the parameter.

If the parameter point $p$ is consistent with the hypothesis $H_0$ to be tested, the probability of making a correct decision is equal to $L(p)$ whereas if the parameter point $p$ is not consistent with the hypothesis $H_0$, the probability of making a correct decision is equal to $1 - L(p)$. Thus, for any parameter point $p$, the probability of making a correct decision can be obtained from the OC function.

An OC function is considered more favourable the higher the value of $L(p)$ for $p$ consistent with $H_0$ and the lower the value of $L(p)$ for $p$ not consistent with $H_0$, that is, when

$$\underset{p \in \Omega_0}{Max\ L(p)} \quad \text{and} \quad \underset{p \in \bar{\Omega}_0}{Min\ L(p)} \tag{9}$$

A sequential test is said to be admissible if for any $p$ in the zone of preference for acceptance, the probability of rejecting $H_0$, $1 - L(p)$, should be less than or equal to a preassigned value $\alpha$, that is,

$$1 - L(p) \leq \alpha \iff L(p) \geq 1 - \alpha \tag{10}$$

and for any $p$ in the zone of preference for rejection, the probability of accepting $H_0$, $L(p)$, should be less than or equal to a preassigned value $\beta$, that is,

$$L(p) \leq \beta \iff 1 - \beta \geq 1 - L(p) \tag{11}$$

which also implies that the test is unbiased. In particular, for an admissible sequential test for the hypotheses (4)

$$L(p_0) = 1 - \alpha \quad \text{and} \quad L(p_1) = \beta \tag{12}$$

and since $\alpha$ and $\beta$ are usually very small numbers then

$$L(p_0) \geq L(p_1) \tag{13}$$

Also for an admissible sequential test

$$\lim_{p \to \infty} L(p) = 0 \quad \text{and} \quad \lim_{p \to 0} L(p) = 1 \tag{14}$$

The Average Sample Number (ASN) function represents the price we have to pay in terms of the number of observations required for the test. The number of observations required by a sequential test is not predetermined, but is a random variable because at any stage of the experiment the decision to terminate the process depends on the results of the observations made so far.

We shall denote by n the number of observations (units) required by the sequential test. Then $n$ is a random variable. Carrying out the same sequential test repeatedly, we shall obtain, in general, different values of $n$. Of particular interest is the expected value of $n$ in the long run, when the same test procedure is applied repeatedly.

For any given test procedure the expected value of $n$ depends only on the distribution of $X$. Since the distribution of $X$ is determined by the parameter point $p$, the expected value of $n$ depends only on $p$. For any given parameter point $p$, we shall denote the expected value of $n$ by $E_p(n)$ and we shall refer to it as the ASN function.

According to Oakland (1950), the formulae for $p, L(p)$ and $E_p(n)$ are given by

$$p = \frac{1 - \left(\frac{q_0}{q_1}\right)^u}{\left(\frac{p_1 q_0}{p_0 q_1}\right)^u - 1}, \tag{15}$$

$$L(p) = \frac{\left(\frac{1-\beta}{\alpha}\right)^u - 1}{\left(\frac{1-\beta}{\alpha}\right)^u - \left(\frac{\beta}{1-\alpha}\right)^u} \tag{16}$$

and

$$E_p(n) = \frac{c_1 + (c_0 - c_1)L(p)}{kp - s} \tag{17}$$

where $u$ is a dummy variable.

For any arbitrarily chosen value $u$, $(p, L(p))$ is a point on the OC curve and $(p, E_p(n))$ is a point on the ASN curve. The OC and the ASN curves can be drawn by plotting a sufficiently large number of points $(p, L(p))$ and $(p, E_p(n))$ respectively corresponding to various values of $u$.

## 3. DISCUSSION ON $L(p)$, $E_p(n)$ AND $D(y)$.

We observe that for an admissible sequential test, using (12), (13) and (14) we have

$$E_{p_0}(n) = \frac{c_1 + (c_0 - c_1)(1-\alpha)}{kp_0 - s},$$ (18

$$E_{p_1}(n) = \frac{c_1 + (c_0 - c_1)\beta}{kp_1 - s},$$ (19

$$E_{p_0}(n) > E_{p_1}(n),$$ (2

$$\lim_{p \to \infty} E_p(n) = 0$$ (2

and

$$\lim_{p \to 0} E_p(n) = -\frac{c_0}{s}$$ (2

Implying $c_0 < 0$ since $s > 0$.

We also observe that **(17)** can be expressed as

$$E_p(n) = \frac{c_1 + (c_0 - c_1)L(p)}{k(p - p^*)}, \qquad p^* = \frac{\log\left(\frac{q_1}{q_0}\right)}{\log\left(\frac{p_1 q_0}{p_0 q_1}\right)}$$ (

implying $E_p(n)$ has a maximum when $p \approx p^*$.

From the OC function $L(p)$, we require more samples when $L(p) = 0.5$, that is, probability that the sequential process will terminate with the acceptance of $H_0$ is the sa as rejecting it when $p$ is the true value of the parameter. Thus, it follows that $L(p^*) =$ and $p^*$ lies in the state of profound indifference. This means the probability of spraying not spraying coincide at 0.5. This is plausible since it is a state of profound indifference hence, most samples need to be taken so that the sequential process eventually terminate either spraying or not spraying.

Considering decision (7), we observe that

$$\frac{dy}{dn} = s = kp^* = (kp)|_{p=p^*},$$ (

that is, the slope of the decision lines is pest density at the state of profound indifference.

## 4. ILLUSTRATION

We shall use the data in Table 1 generated by a mixture of Poisson and gar distributions where gamma distribution is the Pearson Type III with parameters $k = 10$ $p = 12$.

| x | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
|---|---|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|
| $n_x$ | 0 | 3 | 8 | 10 | 2 | 11 | 4 | 4 | 0 | 1 | 2 | 4 | 0 | 1 | 0 | 0 | 0 |

*Source: Maritz and Lwin, 1989*

**Table 1. Mixture of Poisson and Pearson Type III Data**

where $n_x$ is the frequency of $x$.

The mixture distribution is the negative binomial (Kipchirchir, 2011) with parameters $k = 10$ and $p = \frac{1}{2}$. As per the following seven categories:

$$x = \overbrace{0,1,2,}^{1} \; \overbrace{3,}^{2} \; \overbrace{4,}^{3} \; \overbrace{5,}^{4} \; \overbrace{6.}^{5} \; \overbrace{7,8,}^{6} \; \overbrace{9,10,11,12,13,14,15,16}^{7}$$

the data in Table 1 fit negative binomial distribution with parameters $k = 10$ and $p = \frac{1}{2}$ at 5% level of significance. For these data, $\bar{x} = 5$, and the maximum likelihood estimate of $k$ is $\hat{k} = 6.55$.

Using the data in Table 1 to mimic the underlying distribution of the pest, we hypothesize that the farmer should not spray the crops if the number of pests per unit is below average and spray the crops if it is above average (economic threshold level). Thus, the sequential plan can be applied by considering the following hypotheses

$$H_0 : mean \; number \; of \; pests \; per \; unit \; is \; 3 \; or \; better$$

versus $H_1 :$ *mean number of pests per unit is 6 or worse*

chosen arbitrarily with reference to the estimate of the mean $m$ ($\hat{m} = \bar{x} = 5$). If $H_0$ is accepted the crops shall not be sprayed whereas if $H_0$ is rejected ( $H_1$ accepted) the crops shall be sprayed. The hypotheses can be formulated as

$$H_0 : kp \leq kp_0 = 3$$

versus $H_1 : kp \geq kp_1 = 6$

and using maximum likelihood estimate of $k$, we obtain

$$H_0 : p \leq p_0 = \frac{3}{6.55} = 0.458$$

versus $H_1 : p \geq p_1 = \frac{6}{6.55} = 0.916.$

For these hypotheses,

$$s = 4.2589, \quad p^* = 0.65 \quad and \quad p = \frac{1 - (0.761)^u}{(1.522)^u - 1}$$

### 4.1. COMPARING SEQUENTIAL TESTS OF STRENGTHS $(\alpha, \beta) = (0.05, 0.05)$ AND $(\alpha, \beta) = (0.01, 0.01)$

$(\alpha, \beta) = (0.05, 0.05)$ imply both errors are equally not serious whereas $(\alpha, \beta) = (0.01, 0.01)$ imply both errors are equally serious.

Now, for $(\alpha, \beta) = (0.05, 0.05)$:

$$D(y) = \begin{cases} reject\ H_0, & if\ \ y > 4.2589n + 7.011 \\ continue\ sampling, & if\ \ 4.2589n - 7.011 < y < 4.2589n + 7.011 \\ accept\ H_0, & if\ \ y < 4.2589n - 7.011 \end{cases}$$

$$L(p) = \frac{(19)^u - 1}{(19)^u - (0.0526)^u} \quad \text{and} \quad E_p(n) = \frac{7.011 - (14.022)L(p)}{(6.55)p - 4.2589}$$

Next, for $(\alpha, \beta) = (0.01, 0.01)$:

$$D(y) = \begin{cases} reject\ H_0, & if\ \ y > 4.2589n + 10.9413 \\ continue\ sampling, & if\ \ 4.2589n - 10.9413 < y < 4.2589n + 10.9413 \\ accept\ H_0, & if\ \ y < 4.2589n - 10.9413 \end{cases}$$

$$L(p) = \frac{(99)^u - 1}{(99)^u - (0.0101)^u} \quad \text{and} \quad E_p(n) = \frac{10.9413 - (21.8826)L(p)}{(6.55)p - 4.2589}$$
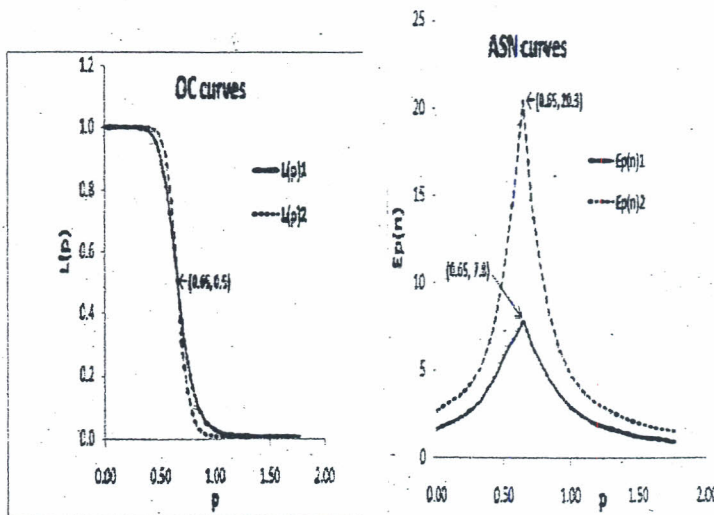


Figure 1. $L(p)1, E_p(n)1$ refers to $(\alpha, \beta) = (0.05, 0.05)$ and $L(p)2, E_p(n)2$ refers to $(\alpha, \beta) = (0.01, 0.01)$

225

The OC curves meet at $p^* = 0.65$ and the OC curve for $(\alpha, \beta) = (0.01, 0.01)$ is greater than the OC curve for $(\alpha, \beta) = (0.05, 0.05)$ when $p < p^*$ and the OC curve for $(\alpha, \beta) = (0.01, 0.01)$ is less than the OC curve for $(\alpha, \beta) = (0.05, 0.05)$ when $p > p^*$. In particular, the OC curve for $(\alpha, \beta) = (0.01, 0.01)$ is greater than the OC curve for $(\alpha, \beta) = (0.05, 0.05)$ in the zone of preference for acceptance of $H_0$ ($p \leq p_0$) and the OC curve for $(\alpha, \beta) = (0.01, 0.01)$ is less than the OC curve for $(\alpha, \beta) = (0.05, 0.05)$ in the zone of preference for rejection of $H_0$ ($p \geq p_1$). Thus, the test for $(\alpha, \beta) = (0.01, 0.01)$ provides better protection against wrong decisions than the test for $(\alpha, \beta) = (0.05, 0.05)$, however, the ASN curve for $(\alpha, \beta) = (0.01, 0.01)$ is greater than for $(\alpha, \beta) = (0.05, 0.05)$, that is, we shall pay a bigger price in terms of observations and more so at the state of profound indifference. In other words, to have better protection against both errors, the sequential process requires more observations and the highest number needed at the state of profound indifference.

The ASN curves have maxima at $p^* = 0.65$ and $L(p^*) = 0.5$ as expected since this is a state of profound indifference. In particular, the test for $(\alpha, \beta) = (0.05, 0.05)$ has a maximum of 8 observations whereas the test for $(\alpha, \beta) = (0.01, 0.01)$ has a maximum of 20 observations (more than double).

### 4.2. COMPARING SEQUENTIAL TESTS OF STRENGTHS $(\alpha, \beta) = (0.01, 0.05)$ AND $(\alpha, \beta) = (0.05, 0.01)$

$(\alpha, \beta) = (0.01, 0.05)$ imply that type I error is more serious than type II error and $(\alpha, \beta) = (0.05, 0.01)$ imply that type II error is more serious than type I error.

Now, for $(\alpha, \beta) = (0.01, 0.05)$ :

$$D(y) = \begin{cases} reject\ H_0, & if\ \ y > 4.2589n + 10.8427 \\ continue\ sampling, & if\ \ 4.2589n - 7.1091 < y < 4.2589n + 10.8427 \\ accept\ H_0, & if\ \ y < 4.2589n - 7.1091 \end{cases}$$

$$L(p) = \frac{(95)^u - 1}{(95)^u - (0.0505)^u} \quad and \quad E_p(n) = \frac{10.8427 - (17.9518)L(p)}{(6.55)p - 4.2589}$$

Next, for $(\alpha, \beta) = (0.05, 0.01)$:

$$D(y) = \begin{cases} reject\ H_0, & if\ \ y > 4.2589n + 7.1091 \\ continue\ sampling, & if\ \ 4.2589n - 10.8487 < y < 4.2589n + 7.1091 \\ accept\ H_0, & if\ \ y < 4.2589n - 10.8487 \end{cases}$$

$$L(p) = \frac{(19.8)^u - 1}{(19.8)^u - (0.0105)^u} \quad and \quad E_p(n) = \frac{7.1091 - (17.98788)L(p)}{(6.55)p - 4.2589}$$
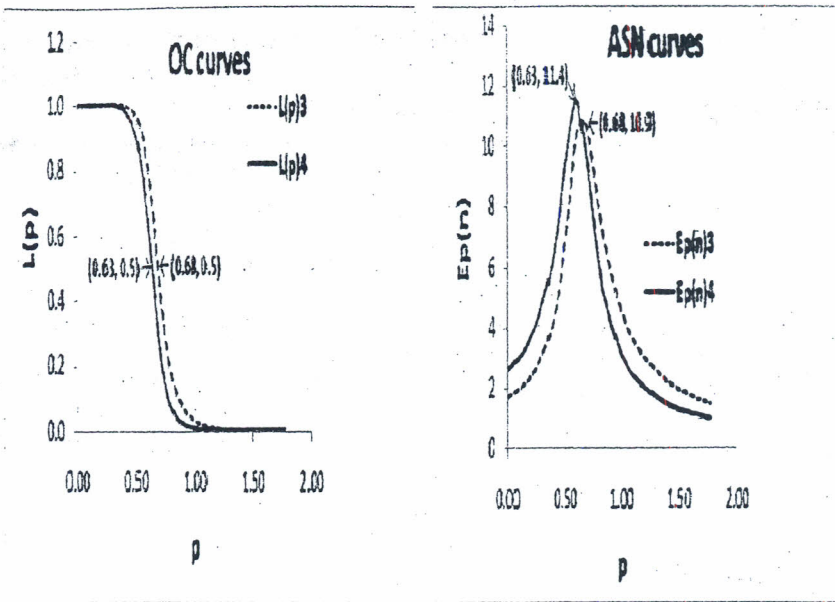
Figure 2. $L(p)3$, $E_p(n)3$ refers to $(\alpha, \beta) = (0.01, 0.05)$ and $L(p)4$, $E_p(n)4$ refers to $(\alpha, \beta) = (0.05, 0.01)$

The OC curve for $(\alpha, \beta) = (0.01, 0.05)$ (type I error is more serious than type II error) is greater than for $(\alpha, \beta) = (0.05, 0.01)$ (type II error is more serious than type I error). Thus, the test for $(\alpha, \beta) = (0.01, 0.05)$ provides better protection against wrong decisions in the zone of preference for acceptance of $H_0$ $(p \leq p_0)$ and the test for $(\alpha, \beta) = (0.05, 0.01)$ provides better protection against wrong decisions in the zone of preference for rejection of $H_0$ $(p \geq p_1)$. Moreover, the test for $(\alpha, \beta) = (0.01, 0.05)$ requires less observations than for $(\alpha, \beta) = (0.05, 0.01)$ in the zone of preference for acceptance of $H_0$ while the test for $(\alpha, \beta) = (0.05, 0.01)$ requires less observations than the test for $(\alpha, \beta) = (0.01, 0.05)$ in the zone of preference for rejection of $H_0$. Thus, the tests provide better protection against wrong decisions in the respective zones without paying a price of taking more observations.

The ASN curves meet at $p^* = 0.65$ and the states of profound indifference, that is, $L(p) = 0.5$, occur at $p = 0.68$ for the test for $(\alpha, \beta) = (0.01, 0.05)$ and at $p = 0.63$ for the test for $(\alpha, \beta) = (0.05, 0.01)$. Thus, states of profound indifference occur at $p \approx p^*$ and both tests have maximum number of 11 observations (more or less half the maximum number of observations of keeping both probabilities of type I and II errors at possible minima).

227

## CONCLUSIONS

From the ASN curves, we observe that for low or high infestation, fewer observations are needed than for moderate infestation. This is plausible since moderate infestation falls in the zone of indifference as depicted by the OC curves and hence more observations need to be taken so that the sequential process eventually terminates at either low or high infestation. Consequently, low infestation will naturally lead to not spraying while high infestation will naturally lead to spraying. If environmental pollution and food security are both serious, we should keep probabilities of both errors at possible minima and have better protection against wrong decisions, however, we pay a price in terms of taking more observations and more so at the state of profound indifference. If we want to spray when necessary, that is, when the infestation has reached economic threshold, we should keep probability of type II error at a possible minimum and if we want to avoid unnecessary pollution of the environment, we should keep probability of type I error at a possible minimum and have better protection against wrong decisions in each case without paying a price in terms of taking more observations.

## ACKNOWLEDGEMENT

## REFERENCES

1. Anscombe, F. J. The statistical analysis of insect counts based on the negative binomial distribution; Biometrics; 5, 1949, pp. 165-174.
2. Ba-Angood, S. A., Stewart, R. K. Sequential sampling for cereal aphids on barley in Southwestern Quebec; Journal of Economic Entomology; 73, 1980, pp. 679-681.
3. Maritz, J. S., Lwin, T. Empirical Bayes methods; Chapman and Hall, 1989.
4. Kipchirchir, I. C. The versatility of the negative binomial distribution in describing dispersion; ICASTOR Journal of Mathematical Sciences; 5(1), 2011, pp. 65-78.
5. Oakland, G. B. An application of sequential analysis to whitefish sampling; Biometrics; 6, 1950, pp. 59-67.
6. Southwood, T. R. E. Ecological Methods; Methuen and Co. Ltd., 1966.
7. Wald, A. Sequential Analysis; John Wiley and Sons Inc., 1947.